

The Cellular Relay Carpet: Distributed Cooperation with Ubiquitous Relaying

Raphael T. L. Rolny · Tim Rüegg · Marc Kuhn · Armin Wittneben

Received: 17 January 2014 / Accepted: 29 May 2014 / Published online: 14 June 2014
© Springer Science+Business Media New York 2014

Abstract We consider the up- as well as downlink of a cellular network in which base stations (BSs) are supported by a large amount of relays spread over the entire area like a carpet. The BSs only see the static relays as the nodes they communicate with, which enables large antenna arrays at the BSs with sophisticated multi-user MIMO transmission. Together with a simple form of BS cooperation, the communication via the small relay cells allows to improve the data rates by distributed interference management and to reduce the complexity at the terminals. We investigate different types of relays as well as different relaying strategies for this relay carpet and compare them with respect to complexity, required channel state information (CSI), and performance in the interference-limited environment of dense cellular networks. The robustness of the different schemes with respect to channel estimation errors is studied and we conclude that especially relays of very low complexity are not sensitive to CSI imperfections. Relays can thus be applied in large numbers and enable massive MIMO at the BSs. The relay carpet proves thereby

to be an efficient approach to enhance future generations of cellular networks significantly.

Keywords Cellular networks · Cooperation · Relaying · Multi-user MIMO · Channel estimation · Imperfections

1 Introduction

The growing demand for ubiquitous data service has led to an ongoing increase of expectations for future cellular networks. The next generations should not only provide data rates that are higher by orders of magnitude than today's systems, but also improved coverage and reliability [2]. In order to stretch the boundaries of cellular systems, spectral efficiency has to be increased. This can be achieved by expanding the networks in the spatial domain, i.e. to introduce more antennas, either physically or virtually. In the former case, the base stations (BSs) can be equipped with (very) large antenna arrays, eventually leading to massive MIMO [3]. Such large arrays allow to serve many users at the same time, for instance using multi-user MIMO methods, and to mitigate the interference in adjacent cells.

An alternative is to increase the BS density and to reduce the cell sizes such that the network consists of pico- or femto-cells [4]. Such small cells can also coexist with micro- or macro-cells in heterogeneous networks [5, 6]. The fundamental advantage is that the adaptation to the user position can be achieved by handovers between cells or sectors, which is easy to implement and requires little overhead. In practice, however, this approach is, among others, limited by the difficulty to identify new BS sites, e.g. due to social acceptance, availability of backbone access etc., and by the cost of deployment. Besides this, it

This paper is partially based on a conference publication [1] presented at the IEEE PIMRC 2013 in London, UK.

R. T. L. Rolny (✉) · T. Rüegg · M. Kuhn · A. Wittneben
Communication Technology Laboratory, ETH Zurich,
8092 Zurich, Switzerland
e-mail: rolny@nari.ee.ethz.ch

T. Rüegg
e-mail: rueegg@nari.ee.ethz.ch

M. Kuhn
e-mail: kuhn@nari.ee.ethz.ch

A. Wittneben
e-mail: wittneben@nari.ee.ethz.ch

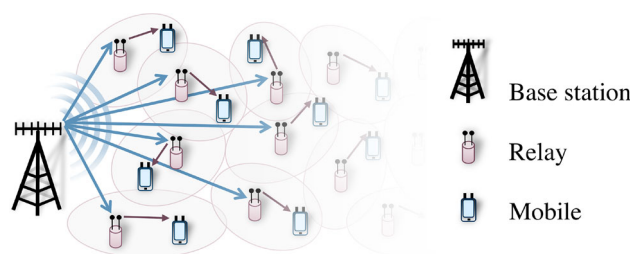


Fig. 1 The relay carpet: a sophisticated BS serves a large amount of MSs in the same physical channel by the help of many distributed relays

is also foreseen that future networks are supported by remote radio units [7] or by wireless relays [8]. These, however, are so far primarily intended for range extension.

Either of these approaches can increase the total throughput of the entire network, while individual user rates remain limited when the mobile stations (MSs) do not have more antennas. As much of the available resources (e.g. bandwidth) as possible should therefore be allocated to each user, up to a reuse factor of one. Due to the interference-limited nature of cellular networks, this can only be achieved by efficient interference management. To this end, current research focuses on BS cooperation/coordinated multipoint (CoMP) transmission that attempts to overcome the detrimental effects of interference on the exploitation of the degrees of freedom in the network [9, 10].

CoMP, however, suffers from severe challenges and difficulties. BSs that perform joint beamforming require very high backhaul rates, not only to support the data rates of their users, but also to exchange user data and channel state information (CSI) with their cooperation partners. Especially if BSs with large arrays are considered, the number of channel coefficients that need to be estimated grows rapidly with the number of involved antennas. This leads to an increasing overhead, as more pilots have to be included in the signals. Achievable performance gains might therefore stagnate or even decrease [11]. Moreover, even when this overhead can be overcome, the performance of CoMP remains limited by residual interference [12].

1.1 Ubiquitous Relaying

An attempt to combine the advantages of the aforementioned approaches, while avoiding their disadvantages, is to support the BSs by a large amount of relays without connection to the backbone. If the relays are of low cost and low power, they can be installed in massive numbers across the entire area of the network, similar to a carpet (see Fig. 1). In this “relay carpet”, the relays serve MSs in their close vicinity within small relay cells. Thereby, many

more relays can be deployed than MSs have to be served. As a result, different subsets of relays can be activated or deactivated, e.g. by BSs that transmit only to selected relays through beamforming or by more sophisticated scheduling functionalities. In this way, static users are served by the relay cells they are located in or multiple relays can follow the movements of mobile users. If GPS information is available, direction and speed of the users can be predicted to assign the relays appropriately. The potential disadvantage that moving users might require many handovers can thus be mitigated.

As a result, the network is turned into a two-hop network in which the BSs as well as the MSs communicate via relays. The BSs see only the relays as their communication partners. If dedicated relays are mounted at fixed positions, fast fading between them is eliminated and the BSs only have to track quasi static channels. This simplifies the estimation of CSI and allows to equip the BSs with (very) large antenna arrays and to apply sophisticated multi-user MIMO transmission. In order to serve mobile users, only the static relays have to be addressed, which is much simpler than to follow possibly fast moving MSs. Accordingly, the static relays enable massive MIMO at the BSs. The MSs on the other hand see a much simpler network of relays with only few antennas in their close vicinity. The relays thus lead to a more equally distributed signal quality and the users experience less pathloss and better coverage.

Additionally, the relays can shape the (effective) channel between BSs and MSs in a beneficial way. Accordingly, network operators do not have to rely on random properties of the propagation channel, which can result in deep fades or shadowed users, but can achieve much more homogeneous coverage. To this end, the relays can perform simple signal processing tasks that allow for signal amplification or even distributed interference management. As a side-benefit, the angular spread of the effective channel can be increased by allocating multiple relays to one user and MSs can be equipped with more antennas in a compact space. As a large amount of relays is deployed, the complexity of the relay nodes is crucial. Different relay architectures can assist the communication between BSs and MSs in different ways, depending on their available CSI and computational power. The node density and the relay complexity thus lead to a tradeoff in which the performance and the infrastructure costs can be balanced.

1.2 Contribution

In this paper, we discuss different relaying schemes and architectures and show how they can improve the performance of future cellular networks by applying them in the ubiquitous form that leads to the relay carpet. We compare

different approaches with respect to achievable rates and complexity and propose methods to cope with the interference in such networks, e.g. based on relay filtering and a specific form of BS cooperation, that are of comparably low complexity. Due to the interference mitigation, high performance gains can be achieved. Particularly two-way relaying proves to be very beneficial in contrast to rather pessimistic results of prior work (cf. e.g. [13]). Furthermore, we investigate the influence of imperfect CSI on these approaches and show that especially simple relays of low complexity are very robust and thus well suited for massive deployment. Parts of this work have already been published in [1], where a specific example of two-way relaying is applied.

The outline of the paper is as follows: In Sect. 2 we describe the relay carpet and the different relay architectures considered in this work and formulate the resulting system models. In Sect. 3, we derive achievable rates for the up- and downlink that will be used as performance measure in the remainder of the paper. Transmission schemes for the terminal nodes and relays are developed in Sect. 4. Aspects of channel estimation at the different nodes are discussed in Sect. 5 and we also describe their effects on the implementation of the relays and terminals. Extensive simulation results that assess the performance of the relay carpet with and without CSI imperfections are presented in Sect. 6. Section 7 finally concludes the paper.

Notation: In the following, boldface lower- and upper-case characters (\mathbf{a} and \mathbf{A}) denote vectors and matrices of complex values. The operators $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate (Hermitian) transpose, respectively. Expectation, trace, determinant, and null space of a matrix are $\mathbb{E}[\cdot]$, $\text{tr}\{\cdot\}$, $\det\{\cdot\}$, and $\text{null}\{\cdot\}$. The $N \times N$ identity matrix is denoted by \mathbf{I}_N .

2 The Relay Carpet Network

The basic organization of the network is similar to a conventional one with micro- or macro-cells. The area is divided into geographically separated cells, each with one BS that is equipped with a large antenna array and multiple MSs that are served simultaneously. The communication between BS and MSs (downlink) and vice versa (uplink) is assisted by a large amount of relays. Different relays can thereby transmit in different frequency bands such that adjacent relay cells form a reuse pattern. Accordingly, the MSs can be served in different resource blocks and the BSs communicate with MSs by assigning an appropriately chosen set of relays.

We consider the relays as dedicated infrastructure nodes that are spread over the entire cell. As such, they are intentionally mounted at fixed positions, e.g. on lamp posts,

at bus stops, or on the wall of a building, and might therefore have a good connection to the BS. Additionally, these links have a long coherence time and fast fading is eliminated. The MSs, on the other hand, are served by small relay cells. If sufficiently many relays are deployed, shadowing effects can be avoided to a large extent.

The relays not only improve the connectivity for the MSs, but can also apply different signal processing tasks. These depend on the architecture of the relays and can range from simple active scattering [14] up to sophisticated filtering, interference cancellation [15], or decoding and encoding [16]. Different implementations can thereby affect the signal processing and the complexity at the other nodes. In the following, we describe the considered relay architectures.

2.1 Relay Architectures

Relays can be classified as full-duplex or half-duplex [17]. While full-duplex relays can simultaneously transmit and receive, half-duplex relays cannot. For instance, half-duplex nodes may operate in *time-division duplex* (TDD) mode, i.e. each node transmits and receives in different time slots; in *frequency-division duplex* (FDD) systems, nodes can transmit and receive at the same time but use different frequency channels. Furthermore, we apply two different signal processing strategies: the *decode-and-forward* (DF) strategy, which involves decoding of the source transmission at the relays before the re-encoded signals are forwarded, and *amplify-and-forward* (AF) relaying, where the relays forward a linear combination of signals at their receive antennas [17].

The complexity of the relays does not only depend on the relaying strategy but also on further implementation aspects, as for instance receive and transmit filters. We consider two different types of relay implementations: In their simplest form, the relays do not use any special receive or transmit filter; the signal is only scaled with a gain matrix given by a scaled identity matrix. We refer to these relays as *type A* relays. The more complex *type B* relays use spatial receive and transmit filters. An especially simple class of relays is given by type A AF relays in an FDD system; such relays can be implemented by a frequency conversion of the received signal. They are not only of very low complexity, but also introduce no (or very small) delays, as the signals are immediately retransmitted. This is not the case for DF relays. Due to the decoding and encoding, the retransmission is delayed by at least a block length, even in FDD mode. Additionally, the DF relays also require the most complex implementation, not only because of the decoding and encoding functionality but also due to the required receive CSI that has to be obtained.

We apply different bidirectional relaying protocols that do not use a direct link between BSs and MSs. These protocols can be classified into *one-way* (conventional) and *two-way* relaying [18]. In the former case, the up- and downlink are separated and the relays either forward the BS signals to the MSs or vice versa. In two-way relaying, both directions of communication are combined such that the relays receive the superposition of all BS and MS signals and broadcast a processed version of these signals back to all terminals. This can double the spectral efficiency as compared to one-way relaying. An inherent drawback of two-way relaying is that the signal received by a terminal (BS or MS) also contains the signal that this terminal has previously transmitted and is backscattered by the relays [19]. This so-called *self-interference* needs to be subtracted at the terminal before the signal can be decoded.

In the remainder of this paper, we study the applicability and performance of relaying for the relay carpet scenario. To this end, we limit ourselves to a preselection of relaying schemes that seem particularly interesting. These schemes arise from the different combinations of the aforementioned aspects (i) duplex mode (TDD/FDD), (ii) relaying strategy (AF/DF), (iii) implementation (type A/B), and (iv) protocol (one-way/ two-way relaying). Thereby, we apply simple linear precoding techniques and discuss the resulting complexity of the different nodes.

2.2 System Model

The network under consideration consists of C cells, each with one BS and multiple MSs. For notational simplicity, we assume that all cells have the same number M of active MSs and that all nodes of the same kind have the same number of antennas, although an extension to a more general case is straightforward. The number of antennas at the BSs is denoted by N_B , the one of the MSs by N_M . The considered communication is bidirectional, i.e. BS c , with $c \in \{1, \dots, C\}$, wants to transmit $d_s \leq N_M$ data streams to MS (c, k) (the k th MS in cell c) in the downlink and, in turn, each MS wishes to send d_s data streams to its BS in the uplink.

As each BS simultaneously serves multiple MSs located in its corresponding cell, we assume $N_B \geq M \cdot N_M$ and write the downlink signal of BS c as

$$\mathbf{x}_c^{(B)} = \sum_{k=1}^M \mathbf{Q}_{c,k}^{(B)} \cdot \mathbf{s}_{c,k}^{(B)}, \quad (1)$$

where $\mathbf{s}_{c,k}^{(B)} \in \mathbb{C}^{d_s}$ is the transmit symbol vector from BS c intended for MS (c, k) and $\mathbf{Q}_{c,k}^{(B)} \in \mathbb{C}^{N_B \times d_s}$ the precoding matrix. In the uplink, the MSs transmit

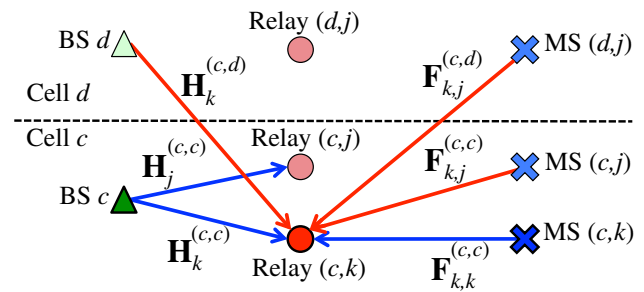


Fig. 2 Network model. The communication between BSs and MSs is assisted by relays

$$\mathbf{x}_{c,k}^{(M)} = \mathbf{Q}_{c,k}^{(M)} \cdot \mathbf{s}_{c,k}^{(M)}, \quad (2)$$

with $\mathbf{s}_{c,k}^{(M)} \in \mathbb{C}^{d_s}$ and $\mathbf{Q}_{c,k}^{(M)} \in \mathbb{C}^{N_M \times d_s}$ being the transmit symbol vector and the precoding matrix of the signal from MS (c, k) intended for BS c .

The bidirectional communication between BSs and MSs is assisted by $K \geq M$ relays. In this paper, we focus on a single resource block, i.e. all relays transmit in the same frequency band. Furthermore, each active MS is served by at least one relay and a relay cannot serve more than one MS.¹ The relays are equipped with N_R antennas, where $N_B \geq N_R \geq N_M$. A sketch of the network can be seen in Fig. 2. The narrow-band channel from BS d to relay (c, k) is denoted by $\mathbf{H}_k^{(c,d)} \in \mathbb{C}^{N_R \times N_B}$ and the reverse channel from relay (c, k) to BS d by $\bar{\mathbf{H}}_k^{(d,c)} \in \mathbb{C}^{N_B \times N_R}$. The channels from MS (d, j) to relay (c, k) and vice versa are denoted by $\mathbf{F}_{k,j}^{(c,d)} \in \mathbb{C}^{N_R \times N_M}$ and $\bar{\mathbf{F}}_{j,k}^{(d,c)} \in \mathbb{C}^{N_M \times N_R}$, respectively. When a TDD protocol is applied, the channels are assumed to be reciprocal, i.e. $\bar{\mathbf{H}}_k^{(d,c)} = \mathbf{H}_k^{(c,d)\top}$ and $\bar{\mathbf{F}}_{j,k}^{(d,c)} = \mathbf{F}_{k,j}^{(c,d)\top}$. If the system is operated in the FDD mode, the channels on the different directions are assumed to be independent. In the following, we describe the end-to-end relations of the system for the different relaying strategies.

2.3 AF One-Way Relaying

In one-way relaying, the up- and downlink are separated, either by different time slots (TDD) or orthogonal frequency bands (FDD). Considering the downlink, the BSs simultaneously transmit their signal (1) and relay (c, k) receives (in the forward direction)

$$\vec{\mathbf{r}}_{c,k} = \sum_{d=1}^C \mathbf{H}_k^{(c,d)} \cdot \sum_{j=1}^M \mathbf{Q}_{d,j}^{(B)} \cdot \mathbf{s}_{d,j}^{(B)} + \vec{\mathbf{n}}_{c,k}, \quad (3)$$

where $\vec{\mathbf{n}}_{c,k}$ is the noise induced in the relay. Assuming AF relaying, the relays multiply their receive signals (3) with a gain matrix $\mathbf{G}_{c,k} \in \mathbb{C}^{N_R \times N_R}$ and, after a possible frequency

¹ More MSs can be served in different frequency bands or by sharing the resources with a TDMA or FDMA scheme.

conversion in FDD, retransmit $\vec{\mathbf{t}}_{c,k} = \mathbf{G}_{c,k} \cdot \vec{\mathbf{r}}_{c,k}$ to the MSs. The receive signal of MS (c, k) is then

$$\vec{\mathbf{y}}_{c,k} = \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{F}}_{kj}^{(c,d)} \cdot \mathbf{G}_{d,j} \cdot \vec{\mathbf{r}}_{d,j} + \mathbf{w}_{c,k}^{(M)}, \quad (4)$$

with $\mathbf{w}_{c,k}^{(M)}$ being the noise in the MS.

In the uplink, the MSs transmit their signals (2) and the (reverse) receive signal at the relays is written as

$$\overleftarrow{\mathbf{r}}_{c,k} = \sum_{d=1}^C \sum_{j=1}^M \mathbf{F}_{kj}^{(c,d)} \cdot \mathbf{Q}_{d,j}^{(M)} \cdot \mathbf{s}_{d,j}^{(M)} + \overleftarrow{\mathbf{n}}_{c,k}. \quad (5)$$

After multiplication of $\overleftarrow{\mathbf{r}}_{c,k}$ with $\mathbf{G}_{c,k}$ and forwarding the resulting signal $\overleftarrow{\mathbf{t}}_{c,k} = \mathbf{G}_{c,k} \cdot \overleftarrow{\mathbf{r}}_{c,k}$, BS c receives

$$\overleftarrow{\mathbf{y}}_c = \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{H}}_j^{(c,d)} \cdot \mathbf{G}_{d,j} \cdot \overleftarrow{\mathbf{r}}_{d,j} + \mathbf{w}_c^{(B)}, \quad (6)$$

where $\mathbf{w}_c^{(B)}$ is the BS noise.

2.4 AF Two-Way Relaying

In two-way relaying, both directions of communication are combined and all BSs and MSs transmit their signals (1) and (2) simultaneously. Accordingly, the relays receive the superposition of all these signals

$$\mathbf{r}_{c,k} = \sum_{d=1}^C \sum_{j=1}^M \left(\mathbf{H}_k^{(c,d)} \mathbf{Q}_{d,j}^{(B)} \mathbf{s}_{d,j}^{(B)} + \mathbf{F}_{kj}^{(c,d)} \mathbf{Q}_{d,j}^{(M)} \mathbf{s}_{d,j}^{(M)} \right) + \mathbf{n}_{c,k}. \quad (7)$$

As before, the AF relays multiply their receive signal vector with a gain matrix $\mathbf{G}_{c,k}$ and broadcast the resulting signal back to all terminal nodes. The resulting signals received by BS c and MS (c, k) are thus given by (8) and (9). These signals not only include the desired signal but also contain what the corresponding node has transmitted itself (self-interference) as well as additional interference from the other nodes of the same kind.

$$\begin{aligned} \mathbf{y}_c^{(B)} = & \sum_{d=1}^C \sum_{j=1}^K \sum_{b=1}^C \sum_{i=1}^M \left(\bar{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{H}_j^{(d,b)} \mathbf{Q}_{b,i}^{(B)} \mathbf{s}_{b,i}^{(B)} \right. \\ & \left. + \bar{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{F}_{j,i}^{(d,b)} \mathbf{Q}_{b,i}^{(M)} \mathbf{s}_{b,i}^{(M)} \right) \\ & + \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{n}_{d,j} + \mathbf{w}_c^{(B)} \end{aligned} \quad (8)$$

$$\begin{aligned} \mathbf{y}_{c,k}^{(M)} = & \sum_{d=1}^C \sum_{j=1}^K \sum_{b=1}^C \sum_{i=1}^M \left(\bar{\mathbf{F}}_{kj}^{(c,d)} \mathbf{G}_{d,j} \mathbf{H}_j^{(d,b)} \mathbf{Q}_{b,i}^{(B)} \mathbf{s}_{b,i}^{(B)} \right. \\ & \left. + \bar{\mathbf{F}}_{kj}^{(c,d)} \mathbf{G}_{d,j} \mathbf{F}_{j,i}^{(d,b)} \mathbf{Q}_{b,i}^{(M)} \mathbf{s}_{b,i}^{(M)} \right) \\ & + \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{F}}_{kj}^{(c,d)} \mathbf{G}_{d,j} \mathbf{n}_{d,j} + \mathbf{w}_{c,k}^{(M)} \end{aligned} \quad (9)$$

2.5 DF One-Way Relaying

In contrast to the AF case, DF relays completely decode the signals they receive before they forward them. The receive signal of relay (c, k) in the downlink is the same as in (3). This signal can then be filtered by a receive combining matrix $\mathbf{G}_{c,k}^{(Rx)H}$, which leads to

$$\vec{\mathbf{r}}_{c,k} = \mathbf{G}_{c,k}^{(Rx)H} \cdot \left(\mathbf{H}_k^{(c,c)} \mathbf{Q}_{c,k}^{(B)} \mathbf{s}_{c,k}^{(B)} + \overrightarrow{\mathbf{x}}_{c,k}^{(R,i+n)} \right), \quad (10)$$

where $\overrightarrow{\mathbf{x}}_{c,k}^{(R,i+n)}$ contains all interference and noise terms.

The symbol vector $\mathbf{s}_{c,k}^{(B)}$ is decoded, while $\overrightarrow{\mathbf{x}}_{c,k}^{(R,i+n)}$ is considered as noise. After that, the relays newly encode the data symbols, possibly with a different code book. Finally, the resulting symbols $\tilde{\mathbf{s}}_{c,k}^{(B)}$ are premultiplied by a transmit filter matrix $\mathbf{G}_{c,k}^{(Tx)}$ and forwarded to the MSs. The receive signal of MS (c, k) follows as

$$\vec{\mathbf{y}}_{c,k} = \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{F}}_{kj}^{(c,d)} \cdot \mathbf{G}_{d,j}^{(Tx)} \cdot \tilde{\mathbf{s}}_{d,j}^{(B)} + \mathbf{w}_{c,k}^{(M)}. \quad (11)$$

In the uplink, the relays receive the signals from the MSs. The receive signal at relay (c, k) , after applying the receive filter, is

$$\overleftarrow{\mathbf{r}}_{c,k} = \mathbf{G}_{c,k}^{(Rx)H} \cdot \left(\mathbf{F}_{k,k}^{(c,c)} \mathbf{Q}_{c,k}^{(M)} \mathbf{s}_{c,k}^{(M)} + \overleftarrow{\mathbf{x}}_{c,k}^{(R,i+n)} \right), \quad (12)$$

where $\overleftarrow{\mathbf{x}}_{c,k}^{(R,i+n)}$ contains the relay noise and all MS interference terms. The relay decodes the corresponding MS symbol vector $\mathbf{s}_{c,k}^{(M)}$, encodes it to $\tilde{\mathbf{s}}_{c,k}^{(M)}$, and multiplies it with $\mathbf{G}_{c,k}^{(Tx)}$. After retransmission, BS c receives

$$\overleftarrow{\mathbf{y}}_c = \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j}^{(Tx)} \tilde{\mathbf{s}}_{d,j}^{(M)} + \mathbf{w}_c^{(B)}. \quad (13)$$

2.6 DF Two-Way Relaying

In the case of two-way relaying, the BSs and MSs transmit simultaneously and relay (c, k) receives

$$\tilde{\mathbf{r}}_{c,k} = \mathbf{G}_{c,k}^{(Rx)H} \left(\mathbf{H}_k^{(c,c)} \mathbf{Q}_{c,k}^{(B)} \mathbf{s}_{c,k}^{(B)} + \mathbf{F}_{k,k}^{(c,c)} \mathbf{Q}_{c,k}^{(M)} \mathbf{s}_{c,k}^{(M)} + \overrightarrow{\mathbf{x}}_{c,k}^{(R,i+n)} \right). \quad (14)$$

Now both data symbol vectors $\mathbf{s}_{c,k}^{(B)}$ and $\mathbf{s}_{c,k}^{(M)}$ are desired. These are decoded by successive interference cancellation (SIC) [20]. The relay can then combine the decoded data streams by an XOR operation with zero padding [21]. The combined data symbol vector $\tilde{\mathbf{s}}_{c,k}^{(R)}$ is precoded by $\mathbf{G}_{c,k}^{(Tx)}$ and the resulting signal is broadcasted. BS c and MS (c, k) then receive this signal under interference from the other relays

$$\mathbf{y}_c^{(B)} = \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{H}}_j^{(c,d)} \cdot \mathbf{G}_{d,j}^{(Tx)} \cdot \hat{\mathbf{s}}_{d,j}^{(R)} + \mathbf{w}_c^{(B)} \quad (15)$$

$$\mathbf{y}_{c,k}^{(M)} = \sum_{d=1}^C \sum_{j=1}^K \bar{\mathbf{F}}_{k,j}^{(c,d)} \cdot \mathbf{G}_{d,j}^{(Tx)} \cdot \hat{\mathbf{s}}_{d,j}^{(R)} + \mathbf{w}_{c,k}^{(M)}. \quad (16)$$

When the relay signal is decoded, the terminals can apply another XOR operation with the data bits they have previously transmitted. With this form of self-interference cancellation, the desired signal can be reconstructed at the terminals [21]. In order to decode all signals from relays in their own cell, the BSs can again apply SIC.

3 Achievable Rates

Once precoding and relay gain matrices are chosen, achievable rates for one-way and two-way relaying can be formulated for both directions of communication. It is thereby assumed that the data symbols in the vectors $\mathbf{s}_{c,k}^{(B)}$ and $\mathbf{s}_{c,k}^{(M)}$ are independent and identically distributed (i.i.d) according to $\mathcal{CN}(0, 1)$. The elements of the noise terms in the relays and terminals, $\mathbf{n}_{c,k}$, $\mathbf{w}_c^{(B)}$, and $\mathbf{w}_{c,k}^{(M)}$, are assumed to be i.i.d. $\mathcal{CN}(0, \sigma_n^2)$ and $\mathcal{CN}(0, \sigma_w^2)$, respectively.

3.1 AF Relaying

In one-way relaying, the achievable rate of the downlink transmission from BS c to MS (c, k) is calculated by

$$\bar{R}_{c,k} = \log_2 \det \left\{ \mathbf{I}_{N_M} + \left(\bar{\mathbf{K}}_{c,k}^{(i+n)} \right)^{-1} \cdot \bar{\mathbf{K}}_{c,k}^{(\text{sig})} \right\}, \quad (17)$$

where $\bar{\mathbf{K}}_{c,k}^{(\text{sig})}$ and $\bar{\mathbf{K}}_{c,k}^{(i+n)}$ are covariance matrices of the desired signal and interference plus noise, which are given in Appendix 1.

In the uplink, we assume that the BSs try to jointly decode all signals from the MSs within their corresponding cell. The achievable sum-rate of the uplink at BS c is thus

$$\bar{R}_c = \log_2 \det \left\{ \mathbf{I}_{N_B} + \left(\bar{\mathbf{K}}_c^{(i+n)} \right)^{-1} \cdot \bar{\mathbf{K}}_c^{(\text{sig})} \right\}, \quad (18)$$

where $\bar{\mathbf{K}}_c^{(\text{sig})}$ is the covariance matrix of the desired signal at BS c that now contains the signals from all MSs in cell c . Accordingly, $\bar{\mathbf{K}}_c^{(i+n)}$ contains the noise as well as the signals originated from all other MSs. These matrices are also derived in Appendix 1.

In the two-way case, the receive signals at the terminals additionally contain the signals these nodes have injected

into the network themselves as well as the signals from the other nodes of the same kind. The achievable rate of the downlink

$$R_{c,k}^{(DL)} = \log_2 \det \left\{ \mathbf{I}_{N_M} + \left(\mathbf{K}_{M,c,k}^{(i+n)} + \mathbf{K}_{M,c,k}^{(\text{self})} \right)^{-1} \cdot \mathbf{K}_{M,c,k}^{(\text{sig})} \right\} \quad (19)$$

thus additionally contains the covariance matrix of the self-interference $\mathbf{K}_{M,c,k}^{(\text{self})}$. For the uplink, we distinguish between interference that is caused by the BSs (including self-interference) and remaining interference from the MSs. The achievable sum rate at BS c follows as

$$R_c^{(UL)} = \log_2 \det \left\{ \mathbf{I}_{N_B} + \left(\mathbf{K}_{B,c}^{(i+n)} + \mathbf{K}_{B,c}^{(\text{BSint})} \right)^{-1} \cdot \mathbf{K}_{B,c}^{(\text{sig})} \right\}, \quad (20)$$

with the covariance matrices given in the Appendix.

3.2 DF Relaying

Achievable rates are also derived for the case when DF relays are used. The one-way case is considered first. When the BSs have transmitted their signals in the downlink and relay (c, k) has applied its receive filter, it decodes the symbol vector $\mathbf{s}_{c,k}^{(B)}$ that is contained in the receive signal (10). To this end, the interference in $\bar{\mathbf{x}}_{c,k}^{(R,i+n)}$ is treated as noise and the resulting rate on the BS-to-relay link can be given as

$$\bar{R}_{c,k}^{(BR)} = \log_2 \det \left\{ \mathbf{I}_{N_R} + \left(\bar{\mathbf{K}}_{R,c,k}^{(i+n)} \right)^{-1} \cdot \bar{\mathbf{K}}_{R,c,k}^{(\text{sig})} \right\} \quad (21)$$

with

$$\bar{\mathbf{K}}_{R,c,k}^{(\text{sig})} = \mathbf{G}_{c,k}^{(Rx)H} \cdot \mathbf{H}_k^{(c,c)} \mathbf{Q}_{c,k}^{(B)} \mathbf{Q}_{c,k}^{(B)H} \mathbf{H}_k^{(c,c)} \mathbf{G}_{c,k}^{(Rx)} \quad (22)$$

and

$$\bar{\mathbf{K}}_{R,c,k}^{(i+n)} = \mathbb{E} \left[\bar{\mathbf{x}}_{c,k}^{(R,i+n)} \cdot \bar{\mathbf{x}}_{c,k}^{(R,i+n)H} \right]. \quad (23)$$

The newly encoded data symbols $\hat{\mathbf{s}}_{c,k}^{(B)}$ are multiplied with $\mathbf{G}_{c,k}^{(Tx)}$ and forwarded to the MSs. The achievable rate on the second hop can similarly be calculated and results in $\bar{R}_{c,k}^{(RM)}$. Finally, an achievable rate of the two-hop link between BS and MS follows as [22]

$$\bar{R}_{c,k} = \min \left\{ \bar{R}_{c,k}^{(BR)}, \bar{R}_{c,k}^{(RM)} \right\}. \quad (24)$$

The end-to-end rate of the uplink can be obtained in a similar way. When the rate of the transmission from MS (c, k) to relay (c, k) is $\bar{R}_{c,k}^{(MR)}$ and the one of the link from this relay to BS c is $\bar{R}_{c,k}^{(RB)}$, the resulting sum rate of the uplink to BS c is

$$\overleftarrow{R}_c = \sum_{k=1}^M \min \left\{ \overleftarrow{R}_{c,k}^{(\text{MR})}, \overleftarrow{R}_{c,k}^{(\text{RB})} \right\}. \quad (25)$$

Note that the achievable rates follow from the assumption that equally long time slots and the same bandwidth is used for both hops. The end-to-end rate could be improved by optimizing the time and frequency allocation of the two individual links. This, however, is unpractical in the cellular context as the different links of the up- and downlink in adjacent cells would not necessarily be separated anymore.

In the case of two-way relaying, both data symbol vectors $\mathbf{s}_{c,k}^{(\text{B})}$ and $\mathbf{s}_{c,k}^{(\text{M})}$ can be decoded using SIC, which leads to a pair of resulting achievable rates $R_{c,k}^{(\text{BR})}$ and $R_{c,k}^{(\text{MR})}$, one for the signal from BS c intended for MS (c, k) and vice versa. The combined and newly encoded relay symbol vector $\tilde{\mathbf{s}}_{c,k}^{(\text{R})} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is precoded and forwarded. BS c and MS (c, k) then receive this signal under interference from the other relays. With self-interference cancellation, the desired signal can be reconstructed at the terminals, where the BSs again apply SIC to decode all signals from their relays. When the achievable rate pairs of the broadcast phase with respect to each relay/user are given by $R_{c,k}^{(\text{RB})}$ and $R_{c,k}^{(\text{RM})}$, the resulting rates achievable on the two-hop up- and downlink are finally

$$R_{c,k}^{(\text{DL})} = \min \left\{ R_{c,k}^{(\text{BR})}, R_{c,k}^{(\text{RM})} \right\} \quad (26)$$

$$R_c^{(\text{UL})} = \sum_{k=1}^M \min \left\{ R_{c,k}^{(\text{MR})}, R_{c,k}^{(\text{RB})} \right\}. \quad (27)$$

Choosing the rates like this ensures that they lie inside the achievable rate region [21]. However, no optimality is claimed. Note that the considered DF scheme requires the relays to decode the complete transmission blocks from both terminals before they can be newly encoded and retransmitted. This introduces additional delays, which could be reduced e.g. with block-Markov coding [22]. This is however not considered here.

3.3 Prelog Factor

Note that we have dropped the prelog factors in the achievable rates, which would occur by multiple channel uses. These factors (e.g. 1, $\frac{1}{2}$, or $\frac{1}{4}$) depend on the specific relaying protocol and the considered duplex mode. Moreover, the relays could convert their BS signals to frequency bands that are currently not used (cf. cognitive radio [23]) or lie in an ISM band. In this case, no additional costs have to be included into the spectral efficiency. The use of

secondary links is especially motivated by the small transmit power of the relays that do not disturb other systems significantly. Moreover, also conventional networks have to divide the resources for the up- and downlink. Thereby, the two directions of communication can use different fractions. For the sake of comparability and to avoid discussions on how the resource blocks are shared, we thus consider two scenarios: (i) in-band relays when the two-way protocol is considered and (ii) a secondary link that is free for the second hop in the case of one-way relaying. In this way, we can omit the prelog factors.

4 Transmission Schemes

In order to gain more understanding in what the limiting factors of the considered network are, we analyze the individual terms of the receive signals at the terminals. To this end, we apply spatially white signaling at all involved nodes and a scaled identity matrix at the relays. In this way, no interference is mitigated and the whole network is flooded with signals. This allows to measure the individual signal contributions for both the up- and downlink and to identify the strongest interference sources. Based on this analysis, we can design precoding and relay gain matrices with the goal to mitigate the most severe interference terms.

We apply a per node transmit power of $P_B = 40$ W at the BSs and $P_R = 6$ W and $P_M = 0.2$ W at the relays and MSs. The precoding and relay gain matrices are accordingly

$$\mathbf{Q}_{c,k}^{(\text{B})} = \sqrt{P_B / (M \cdot N_B)} \cdot \mathbf{I}_{N_B}, \quad \mathbf{Q}_{c,k}^{(\text{M})} = \sqrt{P_M / N_M} \cdot \mathbf{I}_{N_M},$$

and

$$\mathbf{G}_{c,k} = \sqrt{P_R / \text{tr} \left\{ \mathbf{E} \left[\mathbf{r}_{c,k} \cdot \mathbf{r}_{c,k}^H \right] \right\}} \cdot \mathbf{I}_{N_R}.$$

The resulting (averaged) receive signal powers of the BSs and MSs, for one-way as well as two-way relaying are shown in Fig. 3. The network consists of $C = 19$ cells, each containing $M = K = 6$ MSs/relays. In the figure, we distinguish which relays have forwarded the different signal contributions (own relay, other in-cell relays, or relays from other cells in the downlink and own relays and relays from other cells in the uplink). More details on the simulation parameters are given in Sect. 6.

From the figure, we can conclude where the different interference contributions come from. In contrast to one-way relaying, additional interference terms appear in two-way relaying: the signals transmitted by the other terminals of the same kind, including self-interference. These signals are not present in one-way relaying because the different directions of communication are separated by orthogonal

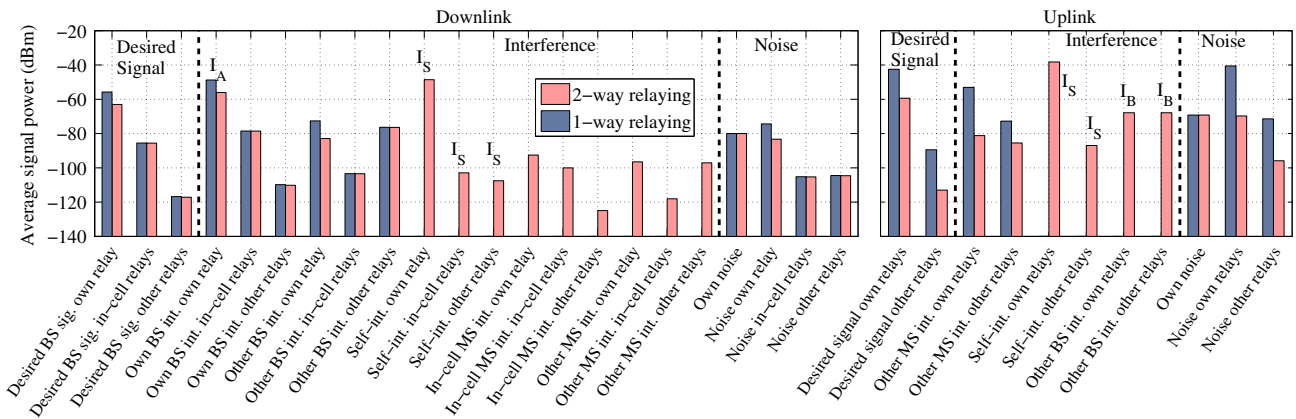


Fig. 3 Receive signal powers distinguished by their sources (one-way and two-way AF relaying protocol)

resources. The total signals received by the relays are thus of less power and one-way relays can apply a higher gain factor in order to meet the transmit power. Consequently, the (existing) signal contributions in one-way relaying are of higher power than in the two-way case. The dominant interference terms can be classified into: (i) BS signals intended for other MSs in the same cell (I_A in the figure), in the case of two-way relaying (ii) self-interference (I_S), (iii) interference from other BSs in the uplink (I_B), and (iv) remaining interference. The terms I_S and I_B do not exist in one-way relaying.

In the following, we apply precoding and relay gain matrices that attempt to mitigate the interference seen by the terminals. A part of the interference can be cancelled with appropriate precoding or beamforming techniques. Other interference terms, however, cannot be reduced without global CSI or cooperation between all nodes. To reduce as much of the interference as possible, we combine different precoding schemes at the different nodes in a way that no sophisticated cooperation between them is required. Thereby, the individual schemes are chosen and combined such that a distributed signal processing for interference mitigation is realized, i.e., the global task of improving the network performance is shared among the different nodes according to their complexity and abilities and each node computes its precoding or gain matrix based on *locally* available CSI. In order to apply schemes that are relevant for practical implementation, we focus on simple linear precoding techniques for which closed form solutions are known and can be computed in a non-iterative fashion. Consequently, we do not claim any optimality of the proposed schemes, but rather understand them as example implementations for the relay carpet that are, due to the low complexity, of high practical relevance. Moreover, the transmissions on the BS-relay links should be independent of the ones on the relay-MS links. This has the advantage that the precoding at the BSs has not to be updated as often as the precoding on the relay-MS links. This is

because the channels between BSs and fixed relays presumably have a much longer coherence time than the channels between the relays and the (possibly moving) MSs. The signaling of the MSs is spatially white such that they do not require any transmit CSI.

4.1 Block Zero-Forcing at the BSs

A strong interference source that degrades the performance in the downlink is the BS signal intended for other MSs (I_A in Fig. 3). To this end, we apply block zero-forcing at the BSs to cancel it [24]. The transmit signal of BS c is

$$\mathbf{x}_c^{(B)} = \sum_{k=1}^M \mathbf{Q}_{c,k}^{(B)} \cdot \mathbf{s}_{c,k}^{(B)} = \sum_{k=1}^M \mathbf{Z}_{c,k} \cdot \tilde{\mathbf{v}}_{c,k} \cdot \mathbf{P}_{c,k} \cdot \mathbf{s}_{c,k}^{(B)}, \quad (28)$$

where

$$\mathbf{Z}_{c,k} = \text{null} \left\{ \left[\mathbf{H}_1^{(c,c)\top}, \dots, \mathbf{H}_{k-1}^{(c,c)\top}, \mathbf{H}_{k+1}^{(c,c)\top}, \dots, \mathbf{H}_M^{(c,c)\top} \right]^\top \right\}$$

ensures that the signal intended for MS (c,k) is nulled at the other relays in this cell and $\tilde{\mathbf{v}}_{c,k}$ are the right hand singular vectors of the virtual channel $\tilde{\mathbf{H}}_k^{(c,c)} = \mathbf{H}_k^{(c,c)} \cdot \mathbf{Z}_{c,k}$. The diagonal power loading matrix $\mathbf{P}_{c,k}$ weights each stream according to the waterfilling solution as in [24].

4.2 AF Relay Gain Matrices

In its simplest form, AF relaying is performed with a scaled identity matrix

$$\mathbf{G}_{c,k} = \sqrt{P_R / \text{tr} \left\{ \mathbf{E} \left[\mathbf{r}_{c,k} \cdot \mathbf{r}_{c,k}^H \right] \right\}} \cdot \mathbf{I}_{N_R}. \quad (29)$$

These type A relays forward their receive signal scaled according to the power constraint, without modifying it. This form of AF relaying does not require any CSI at the relays.

More sophisticated type B relays that have access to local CSI can form linear combinations of all input streams to a beneficial output signal vector. The relay can e.g. design the relay gain matrix such that undesired signals are minimized while the desired signal components should remain at a good quality. To this end, the relay gain matrices are factorized to

$$\mathbf{G}_{c,k} = \sqrt{\alpha_{c,k}} \cdot \mathbf{G}_{c,k}^{(\text{Tx})} \cdot \mathbf{G}_{c,k}^{(\text{Rx})\text{H}}, \quad (30)$$

where $\mathbf{G}_{c,k}^{(\text{Rx})}$ is a receive filter, $\mathbf{G}_{c,k}^{(\text{Tx})}$ a transmit filter, and $\alpha_{c,k}$ a scaling factor to adjust the transmit power.

For the design of the receive filter, we distinguish between one-way and two-way relaying. In the one-way case, the receive filter $\mathbf{G}_{c,k}^{(\text{Rx})}$ is chosen to suppress the interference coming from the BSs of adjacent cells. Such a filter can be obtained by $\mathbf{G}_{c,k}^{(\text{Rx})} = [\mathbf{v}_1^{(c,k)}, \dots, \mathbf{v}_{d_s}^{(c,k)}]$ [25]. Therein, $\mathbf{v}_i^{(c,k)}$ is the eigenvector corresponding to the i th smallest eigenvalue of

$$\mathbf{\Gamma}_{c,k} = \sum_{\substack{d=1 \\ d \neq c}}^C \mathbf{H}_k^{(c,d)} \cdot \mathbf{H}_k^{(c,d)\text{H}}. \quad (31)$$

With this, the receive signal is projected into the subspace that contains the least BS interference under the assumption of spatially white signaling. This has the advantage that $\mathbf{G}_{c,k}^{(\text{Rx})}$ is independent of the actual BS signals and has thus not to be updated when a BS changes its precoding. Moreover, when the relay position is fixed, this covariance matrix is mainly static and simple to estimate.

In two-way relaying, we can additionally enhance the uplink performance by choosing a receive filter that does not only reduce the interference from adjacent BSs but tries also to keep the signal from its MS at a good quality. To this end, $\mathbf{G}_{c,k}^{(\text{Rx})}$ can be chosen as a filter that minimizes the BS interference and noise under the constraint that the MS signal is kept constant. The resulting optimization problem

$$\begin{aligned} \mathbf{G}_{c,k}^{(\text{Rx})} = \arg \min \{ & \mathbf{G}_{c,k}^{(\text{Rx})\text{H}} (\mathbf{\Gamma}_{c,k} + \sigma_n^2 \mathbf{I}_{N_R}) \mathbf{G}_{c,k}^{(\text{Rx})} \} \\ \text{such that } & \mathbf{G}_{c,k}^{(\text{Rx})\text{H}} \mathbf{F}_{k,k}^{(c,c)} = \mathbf{I}_{N_M} \end{aligned} \quad (32)$$

can be solved in closed form and its solution is given by

$$\mathbf{G}_{c,k}^{(\text{Rx})} = (\mathbf{\Gamma}_{c,k} + \sigma_n^2 \mathbf{I}_{N_R})^{-1} \cdot \mathbf{F}_{k,k}^{(c,c)} \cdot (\mathbf{F}_{k,k}^{(c,c)\text{H}} (\mathbf{\Gamma}_{c,k} + \sigma_n^2 \mathbf{I}_{N_R})^{-1} \mathbf{F}_{k,k}^{(c,c)})^{-1}. \quad (33)$$

This approach is a MIMO extension of the minimum variance distortionless response (MVDR) filter [26].

The transmit filter of the relay is chosen as a transmit matched filter (MF)

$$\mathbf{G}_{c,k}^{(\text{Tx})} = \bar{\mathbf{F}}_{k,k}^{(c,c)\text{H}} \quad (34)$$

with respect to the channel to the corresponding MS. The combined relay gain matrix is then scaled with

$$\alpha_{c,k} = \frac{P_R}{\text{tr} \left\{ \mathbf{G}_{c,k}^{(\text{Tx})} \mathbf{G}_{c,k}^{(\text{Rx})\text{H}} \mathbf{E} \left[\mathbf{r}_{c,k} \mathbf{r}_{c,k}^{\text{H}} \right] \mathbf{G}_{c,k}^{(\text{Rx})} \mathbf{G}_{c,k}^{(\text{Tx})\text{H}} \right\}}.$$

Note that the gain matrices of these type B relays are chosen such that the relays mainly improve the links to the MSs, because the BS-relay links are presumably already strong due to the high transmit power and the zero-forcing at the BSs. Also note that the receive filters at the relays depend only on the covariance matrix of the BS-relay interference. The individual channel coefficients need not to be known. Moreover, the relay receive filters do not have to be updated very often, since these channels change only slowly when the relays are at fixed positions. Additionally, the precoding at the BSs can be done with respect to the effective channel that includes the specific relay receive filters, i.e. the block zero-forcing and waterfilling is given as a function of the effective channel $\mathbf{G}_{c,k}^{(\text{Rx})\text{H}} \cdot \mathbf{H}_k^{(c,c)}$ instead of $\mathbf{H}_k^{(c,c)}$ only. This further improves the overall performance.

4.3 DF Relay Filter Design

The same filter techniques can also be applied to DF relays. When type A DF relays are considered, the relay filter matrices are $\mathbf{G}_{c,k}^{(\text{Rx})\text{H}} = \mathbf{I}_{N_R}$ and $\mathbf{G}_{c,k}^{(\text{Tx})} = \sqrt{P_R/N_R} \cdot \mathbf{I}_{N_R}$.

For the more sophisticated type B relays, the filters from the AF case can be adopted. In this case, the receive filter in the one-way protocol contains, as for AF relaying, the eigenvectors corresponding to the d_s smallest eigenvalues of $\mathbf{\Gamma}_{c,k}$, i.e. $\mathbf{G}_{c,k}^{(\text{Rx})} = [\mathbf{v}_1^{(c,k)}, \dots, \mathbf{v}_{d_s}^{(c,k)}]$. This projection not only reduces the BS interference, but also results in a smaller dimension of the (effective) signal space seen by the BSs. As a result, the BSs need to zero-force fewer dimensions and thus have additional antennas to improve their beamforming. The transmit filter is a scaled transmit MF

$$\mathbf{G}_{c,k}^{(\text{Tx})} = \sqrt{\frac{P_R}{\text{tr} \left\{ \bar{\mathbf{F}}_{k,k}^{(c,c)\text{H}} \cdot \bar{\mathbf{F}}_{k,k}^{(c,c)} \right\}}} \bar{\mathbf{F}}_{k,k}^{(c,c)\text{H}}, \quad (35)$$

such that it meets the relay transmit power constraint.

For two-way relaying, the receive filter can be replaced by the MVDR solution as in (33).

4.4 Self- and BS-Interference Cancellation

In two-way relaying, both directions of communication are combined into the same physical channel. A strong contribution of interference is thus the self-interference that

propagates back from the relays (I_S in Fig. 3). This interference, however, can be canceled at each node in case of AF relaying. Thereby, the effective channel from itself via the relays back to it has to be known. At the MSs, this effective channel is described by an $N_M \times N_M$ matrix which can be estimated with pilot symbols that are included in the MS signal. Alternatively or in addition, the self-interference can be used to obtain CSI estimates [19]. When the self-interference is completely cancelled, the covariance matrix $\mathbf{K}_{M,c,k}^{(\text{self})}$ in (19) disappears and the resulting rate is significantly improved.

Self-interference can also be canceled at the BSs in the same way. However, this might not be sufficient to achieve high uplink rates, as the sum of signals from all other BSs $d \neq c$ is a strong contribution of the interference at BS c (I_B in Fig. 3). Therefore, we propose that (at least close) BSs cooperate with each other in a way that they share their transmit symbols. In this way, the BSs can not only cancel their self-interference, but can also reconstruct and cancel the interference caused by neighboring BSs. The known data symbols or pilot/ training sequences included in the signals can be used to estimate the effective channels via the relays and no CSI needs to be shared. As a result, the covariance matrix $\mathbf{K}_{B,c}^{(\text{BSint})}$ disappears in (20) completely. This form of BS cooperation improves the uplink rates of two-way relaying drastically.

4.5 Performance Evaluation

In Fig. 4, we preview achievable rates of the aforementioned transmission schemes compared with the rates of a reference case in which no relays are used and the BSs serve the MSs directly by block zero-forcing. The BSs, relays, and MSs transmit with a fixed transmit power of $P_B = 40$ W, $P_R = 6$ W, and $P_M = 0.2$ W, respectively. As discussed in Sect. 3.3, the prelog factors of the relaying schemes are intentionally omitted for TDD and FDD relaying. For comparison, however, we also include the rates of TDD relaying when this factor that arises from the multiple channel uses for one-way relaying is considered (in-band relays); this factor has no impact on two-way relaying. It can be seen that significant gains can be achieved with the relay carpet, even with the simple type A relays. When the prelog factor is taken into account, two-way relaying leads to the best results, as the up- and downlink are combined in a spectrally efficient way. For one-way relaying, the performance is somewhat diminished. Nevertheless, the use of the relays has still its advantages. The acquisition of CSI is drastically simplified and massive MIMO is enabled at the BSs. The achievable rates of the reference are thus rather optimistic, as the

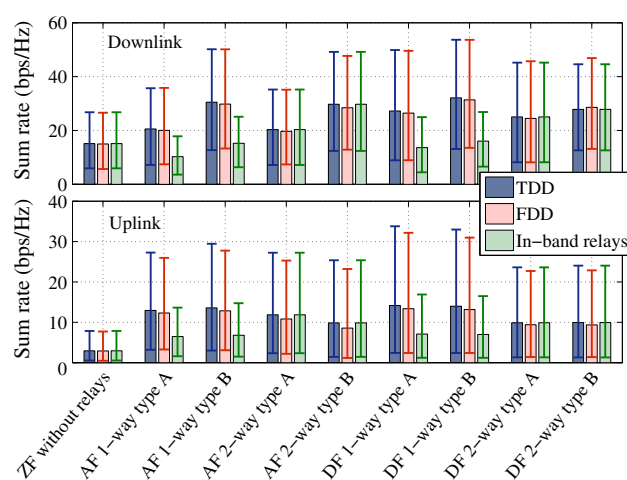


Fig. 4 Achievable sum rates of the different relaying schemes compared to a conventional network without relays in which the BSs serve the MSs directly with zero-forcing. Shown are the mean rates (bars) as well as 5 and 95 % percentiles

overhead to obtain the required CSI, especially from moving MSs, is not considered. The performance of TDD and FDD relaying is comparable. The reciprocal channels in the TDD case do not have a significant impact. For FDD systems, very similar results can be expected. The following simulations are therefore limited to the TDD case and the prelog factor is no longer considered.

5 Aspects of Channel Estimation

For the transmission schemes introduced in the previous section, CSI is necessary at the BSs, relays, and MSs in different forms. We distinguish between CSI at the receiver (CSIR) and at the transmitter (CSIT). Usually, acquiring CSIR (e.g. based on a training sequence) is not considered as difficult as obtaining CSIT. In TDD systems assuming channel reciprocity, CSIT can be determined from the CSIR which has been obtained as part of the decoding process in a previous transmission. In case of FDD this is not possible due to the different frequencies. One way of acquiring CSIT nevertheless is using a feedback channel: the receiver is feeding its CSIR (possibly quantized and compressed) back to the sender. However, CSIT may then be outdated or noisy (e.g. due to quantization). Another way to acquire CSIT in a FDD system would be that the receiver transmits a training sequence on the transmit frequency of the transmitter in a separate time slot. The transmitter then estimates the CSIR and determines the CSIT assuming channel reciprocity. In the following, we discuss which nodes need which form of CSI, how they can acquire it, and what impact this has on the node

complexity. Furthermore, we introduce error models for CSI imperfections in order to determine the robustness of the transmission schemes.

5.1 Acquisition of CSI

At the BSs, channel estimation is necessary for different tasks: CSIR to decode the uplink signals, CSIR to cancel self- and BS-interference and CSIT for the calculation of the beamforming matrices. Whereas CSIR can be acquired at the BS based on training sequences as described above, CSIT needs to be estimated at the relay and fed back to the BS, or the relay can transmit a training sequence on the transmit frequency of the BS on demand.

At the MSs, no CSIT is required by the schemes presented in this paper, only CSIR for decoding the downlink signal and, in case of two-way relaying, for canceling the self-interference.

At the relays, the necessity of CSIR and CSIT depends on the type of the relay and the signal processing. Whereas a type A AF relay does not need any CSI at all, a type B AF relay needs to know the relay-MS channel $\mathbf{F}_{k,k}^{(c,c)}$ as well as the BS interference covariance matrix $\mathbf{\Gamma}_{c,k}$ for the computation of the transmit and receive filter. For TDD relays, the CSIT can be acquired via the CSIR. For FDD, either a feedback from the BS/MS is necessary or the transmission of training sequences by the BS and MS on the transmit frequency of the relay. As only the channel covariance matrix from the BSs is required, the estimation is much simpler than for the full channel. Its dimensions are only $N_R \times N_R$ and a sample covariance matrix can be obtained by observing the received signal over time. For DF relays, CSIR is always necessary for the decoding, also for type A relays. When type B DF relays are used, the CSIT can be obtained from CSIR when the relays operate in a TDD mode. In the FDD case, this is not possible and the acquisition of CSIT by feedback or pilot transmission comes on top.

5.2 Node Functionality

The simplest form of relays considered in this paper are type A AF relays in FDD mode. In this case, the relays can be seen as simple frequency converters that amplify their input signal without the requirement of any CSI. In order to allow its BS to estimate the BS-relay channel, these relays have to be able to transmit a training sequence on demand. This kind of relay can be referred to as a “drilled” relay, as it only responds to requests of the BS. Apart from some synchronization mechanisms, such relays do not need any additional functionalities. If the relays operate in TDD

mode, an additional buffer to store the received signal before it can be retransmitted is required.

The more sophisticated type B AF relays additionally need to acquire CSI such that they can calculate their receive and transmit filters. To this end, the relays need either to be able to estimate the required channels themselves or to receive the CSI that is delivered from their BS and/or MS. As a result, such relays require a decoding functionality that does not differ much from the one in DF relays.

DF relays are the most complex relays considered in this paper. Additional to the CSIR necessary for the decoding, the signals need to be re-encoded. For type B DF relays, also CSIT is required that can be obtained as in the AF case.

While the relaying protocol (whether one-way or two-way) does not matter for the relay complexity in AF relays, it influences the tasks of the terminal nodes. For one-way relaying, the terminal nodes just need to evaluate the training sequences and decode the signal. For two-way relaying instead, they additionally need to estimate and subtract the self-interference (and the interference of the other BSs). Especially for the BSs, that cancel the other BS interference, two-way relaying thus adds some complexity to the terminals. However, when the relays are static, the CSI for interference cancellation needs to be tracked with a comparably low frequency. If DF relays are used, the task of interference cancellation is simpler. Only self-interference has to be compensated, which can be done in the digital domain by an XOR operation.

5.3 Estimation Error Models

As the positions of BSs and relays are fixed, we consider the channel between a BS and a relay as quasi-static. Acquiring CSIT of a certain quality for this link seems possible and less difficult than for the link between a relay and a possibly moving MS. These considerations motivate the chosen transmission schemes.

In the following, we investigate the robustness of the considered schemes regarding imperfect CSI. These imperfections can arise from channel estimation errors, quantization of the channel estimates in the feedback channel, outdated CSI, etc. In order to capture these effects, we apply simple models that are based on additive Gaussian errors as e.g. in [27].

Complete Channel Matrix

For the BS beamforming and the relay filters, the actual channel matrices $\mathbf{H}_k^{(c,c)}$ and $\mathbf{F}_{k,k}^{(c,c)}$ need to be known at the respective nodes. Imperfections on this type of CSI is modeled as

$$\hat{\mathbf{H}}_k^{(c,c)} = \sqrt{L_p} \left(\sqrt{1 - \vartheta_H^2} \mathbf{H}_k^{(c,c)} + \vartheta_H \mathbf{W}_k^{(c,c)} \right), \quad (36)$$

where $\vartheta_H^2 \in [0, 1]$ is the CSI noise scaling factor and the pathloss L_p is assumed to be known perfectly (averaged over time). Only the small scale fading is affected by the estimation error $\mathbf{W}_k^{(c,c)}$ with elements $\mathcal{CN}(0, 1)$. We define the estimation signal-to-noise ratio (SNR) as $\text{SNR}_H = \frac{1 - \vartheta_H^2}{\vartheta_H^2}$ as a measure for the quality of the CSI. As the channels between the BSs and the relays are considered quasi-static, high SNRs can be expected.

For the estimation of $\mathbf{F}_{k,k}^{(c,c)}$, the same model is used. Thereby, the estimation SNR given by $\text{SNR}_F = \frac{1 - \vartheta_F^2}{\vartheta_F^2}$ can differ from the one at the BS, as this channel cannot be assumed to be quasi-static.

Channel Covariance Matrix

For the error of the estimation of the covariance matrix $\mathbf{F}_{c,k}$, required for the calculation of $\mathbf{G}_{c,k}^{(\text{Rx})}$, we use the model

$$\hat{\mathbf{F}}_{c,k} = \mathbf{F}_{c,k} + \sigma_F^2 \mathbf{W}_{c,k} \mathbf{W}_{c,k}^H, \quad (37)$$

where $\mathbf{W}_{c,k}$ is again an estimation error matrix as above and $\sigma_F^2 \in [0, \infty)$ the noise scaling factor. The instantaneous estimation SNR of this model is defined as $\text{SNR}_F = \frac{\text{tr}\{\mathbf{F}_{c,k}\}}{N_R \sigma_F^2}$. The estimation error is assumed to be small, as the sample covariance matrices can be averaged over time.

CSI for Interference Cancellation

For the cancellation of the self-interference at the BSs and the MSs, we consider the compound channels (from the BS/MS to the relays and back) denoted by $\mathbf{H}_{c,k}^{(\text{comp})}$ and $\mathbf{F}_{c,k}^{(\text{comp})}$. These can be estimated e.g. with training sequence contained in the transmit signals. The estimation error of the compound channels is modeled by

$$\hat{\mathbf{H}}_{c,k}^{(\text{comp})} = \mathbf{H}_{c,k}^{(\text{comp})} + \sigma_s \mathbf{W}_{c,k}, \quad (38)$$

with $\mathbf{W}_{c,k}$ the estimation error matrix as above and $\sigma_s \in [0, \infty)$ the CSI noise scaling factor. For the cancellation of the self-interference, the BS/MS subtracts the estimated self-interference. For the achievable rate, only the remainder of the self-interference covariance matrix is of importance

$$\hat{\mathbf{K}}_{M,c,k}^{(\text{self})} = \sigma_s^2 \mathbf{W}_{c,k} \mathbf{W}_{c,k}^H. \quad (39)$$

To relate the estimation noise power to the actual self-interference power, we define the instantaneous estimation SNR of this error model as $\text{SNR}_{\text{self}} = \frac{\text{tr}\{\mathbf{K}_{c,k}^{(\text{self})}\}}{N_i \sigma_s^2}$, for $i \in \{\text{B}, \text{M}\}$.

The same model is used for the cancellation of the interference from other BSs. The remainder of the other BS signal covariance matrix is modeled as

$$\hat{\mathbf{K}}_{B,c}^{(\text{BSint})} = \sigma_B^2 \mathbf{W}_{c,k} \mathbf{W}_{c,k}^H, \quad (40)$$

with all parameters as above. As these channels are assumed to be quasi-static and all data is expected to be known at the receiver, high SNRs can be expected.

6 Simulation Results

We study the performance of the described relay carpet approach by means of computer simulations in a realistic setup. We focus on the sum rate that is achievable in a cell of interest and compare the performance to a non-cooperative reference scenario, that is a cellular network without relays in which the BSs serve multiple MSs by block zero-forcing and waterfilling on the direct BS-MS channels.

Simulation Setup

The network consists of $C = 19$ hexagonal cells, where 18 cells are arranged in two circles around a middle cell that is the cell of interest. The distance between adjacent BSs is 1,000 m. Each cell contains $M = K$ MSs/relays with $N_M = 2$ and $N_R = 4$ antennas. The BS antenna arrays have $N_B = M \cdot N_R$ antennas. All antennas are omnidirectional and we apply the WINNER II channel model as in [28] to get a realistic network model. The channels are drawn according to the WINNER II scenario C2 with line-of-sight condition for all channels between a BS and its associated relays. For all other channels, we impose a non-line-of-sight condition. If not stated otherwise, the chosen transmit powers at the BSs, relays, and MSs are $P_B = 40$ W, $P_R = 6$ W, and $P_M = 0.2$ W. Assuming a total bandwidth of 100 MHz and a noise figure of 5 dB at all nodes, the noise variances are $\sigma_n^2 = \sigma_w^2 = 5 \times 10^{-12}$ W.

Perfect CSI at All Nodes

Empirical cumulative distribution functions (CDFs) of achievable sum rates in the down- as well as uplink are shown in Fig. 5, where $K = M = 6$ relays are placed at a distance of $d_{BR} = 350$ m in a circle around each BS. The MSs are uniformly distributed in small relay cells such that each MS is served by one relay. In a zone of $\frac{2}{3}d_{BR}$ around the BS, no MSs are considered. Users in this area can be served by other relays operating in other frequency bands. Alternatively, static MSs located close to the BS can also be served by the BS directly. By applying such a “deadzone”, we only consider MSs that are located towards the cell edge. Such cell-edge users are particularly challenging in the context of interference-limited cellular networks. The CDFs show that, compared to the non-cooperative reference scenario, very high gains can be achieved by the relay carpet approach. As stated in Sect. 3.3, prelog factors are not considered in the presented achievable rates. Hence, in case all resources have to be counted and the relays are

half-duplex, the rates of all considered variants of one-way relaying must be scaled with $1/2$ before comparing them to the reference scenario or the two-way relaying schemes (cf. Fig. 4). Then, two-way relaying outperforms one-way relaying in all investigated schemes. If the resources of the second hop do not have to be accounted for or full-duplex relays can be used, the one-way curves show their potential gains compared to the two-way approach due to higher gain factors and less interference.

In the following, we only look at the sum of the up- and downlink rates. Figure 6 shows these sums versus the distance between BS and relays for selected type B relays. The results represented by solid lines (with deadzones around the BSs) confirm the results in the CDFs of Fig. 5, which were found for $d_{BR} = 350$ m. With increasing d_{BR} , the performance of the considered relaying schemes improves up to 400 m; only for $d_{BR} < 200$ m the reference scenario performs better. If we compare the rates also for the case without deadzones, shown as dashed lines in Fig. 6, the gains look less impressive. In the reference scenario, MSs that are very close to a BS achieve very high rates by the direct BS transmission with waterfilling, which favors strong users. Additionally, the direct transmission seems to be an aggressive reference because the BSs would have to track the channels to mobile users with many antennas. With the relay carpet, the CSI estimation at the BSs is simplified as fast fading is eliminated from the point of view of the BSs, since the relays are, in contrast to the MSs, not moving. Moreover, the relay schemes achieve much higher rates on the cell edge whereas in the case of direct transmission, the high rates that contribute most to the average are for MSs located very close to the BS. The relaying schemes thus lead to a more balanced and fairer rate distribution.

Achievable sum rates for varying transmit powers are shown in Fig. 7. Here, the distance between BSs and relays is again $d_{BR} = 350$ m and a deadzone is applied. The curves show that, while the network is still interference-limited, notably steeper slopes can be achieved with the relays. This indicates that in the regime around 20–40 dBm, more degrees of freedom can be exploited. Interestingly, AF relaying performs very good, even though this relaying strategy also amplifies noise and interference. Similar performance can only be achieved by DF relaying in the one-way protocol. This type of relaying, however, requires 4 orthogonal resources for one transmission in each direction.

Imperfect CSI

So far, perfect CSI was assumed for all simulations, i.e. the beamforming and relay gain matrices are all computed based on the correct channels. In the following, we study the influence of CSI imperfections as discussed in Sect. 5. The influence of the CSI noise on the up- and downlink

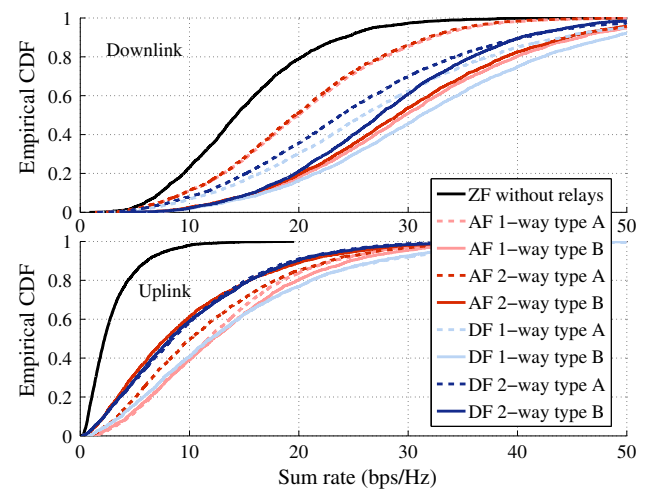


Fig. 5 CDF of achievable sum rates for up- and downlink. The transmit powers are $P_B = 40$ W, $P_R = 6$ W, $P_M = 0.2$ W, and noise variances $\sigma_n^2 = \sigma_w^2 = 5 \times 10^{-12}$ W per 100 MHz bandwidth

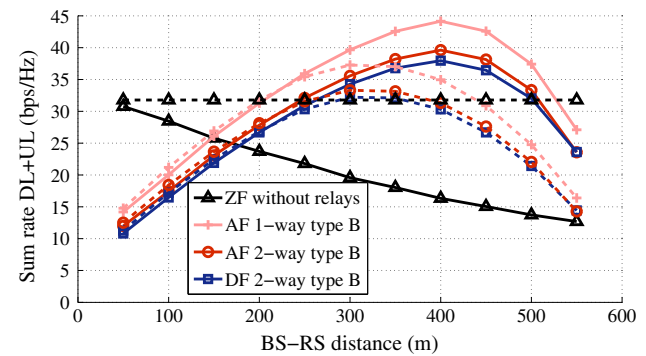


Fig. 6 Average sum rates (up- plus downlink) for different d_{BR} . The solid lines are with a deadzone of $2d_{BR}/3$ around the BSs, the dashed lines are without deadzone

performance is shown in Fig. 8. In the first three columns, only one type of CSI imperfections is considered at one time: (i) only at the BSs for the calculation of the beamforming, (ii) only at the relays, and (iii) only for interference cancellation at the terminals. In the rightmost column, all nodes are affected by CSI imperfections in the same way, i.e. all estimation SNRs are equal.

- (i) It can be seen that the BS beamforming requires good CSI. Otherwise, the performance degrades rapidly. This is not surprising, as zero-forcing is known to be sensitive with respect to channel knowledge. Nonetheless, as we consider the channels between BSs and relays as quasi-static, a high CSI estimation SNR can be expected in our setup. In the uplink, only two-way relaying depends on the BS beamforming.

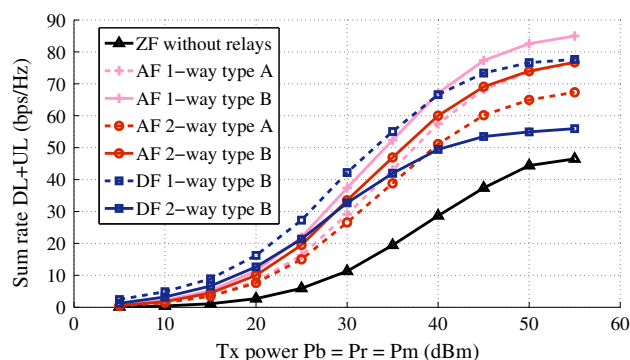


Fig. 7 Achievable sum rates (up- plus downlink) for varying transmit powers. The noise variances are fixed to $\sigma_n^2 = \sigma_w^2 = 5 \times 10^{-12}$ W

- (ii) At the relays, CSI imperfections only have an influence on type B relays. The chosen relaying schemes are however quite robust; the interference mitigating receive filter and the transmit matched filter do not degrade the performance significantly at low SNRs.
- (iii) In the case of AF two-way relaying, the cancellation of self- and BS-interference is crucial. This type of interference is very strong at all nodes and has thus to be known accurately in order to get good end-to-end performance. This form of relaying is thus only beneficial if the terminals can estimate the corresponding channels appropriately, especially at the BSs where the interference from other BSs has also to be cancelled.

Denser Cellular Network

In the previous simulations, all cells contain $K = M = 6$ relays and MSs. However, the gains achievable with the relay carpet can be further increased with more nodes. Figure 9 shows average sum rates for different numbers of users, where $M = K$ and $N_B = M \cdot N_R$ grow accordingly. The relays are randomly placed with a uniform distribution in the cell with a deadzone of 300 m around the BSs. The transmit powers are again $P_B = 40$ W, $P_R = 6$ W, $P_M = 0.2$ W. The curves are plotted for the case of perfect CSI at all nodes (solid lines) as well as for the case in which the different nodes are affected by CSI estimation errors (dashed lines). In the latter case, the BS beamforming is based on CSI with an SNR of 20 dB, the CSI at the relays has an SNR of 10 dB, and the one for the self- and BS-interference cancellation has an SNR of 30 dB. While adding more and more relays into the system, the total transmit power of each cell also increases, as the transmit power of each node is fixed. To this end, a curve of the reference scheme in which the BSs have perfect CSIT and transmit with a power that corresponds to the total power of all nodes in the cell, i.e. $\tilde{P}_B = P_B + K \cdot P_R$, is also included (black dotted line). With this comparison, the performance gains that are due to the higher power can be differentiated from the improvements that come from the relaying schemes.

It can be seen that the performance of AF relaying improves with the number of relays/MSs, while the other schemes tend to saturate with the number of users. This indicates that the impact of interference is weaker with AF relays. Even though a part of the gains of the relaying

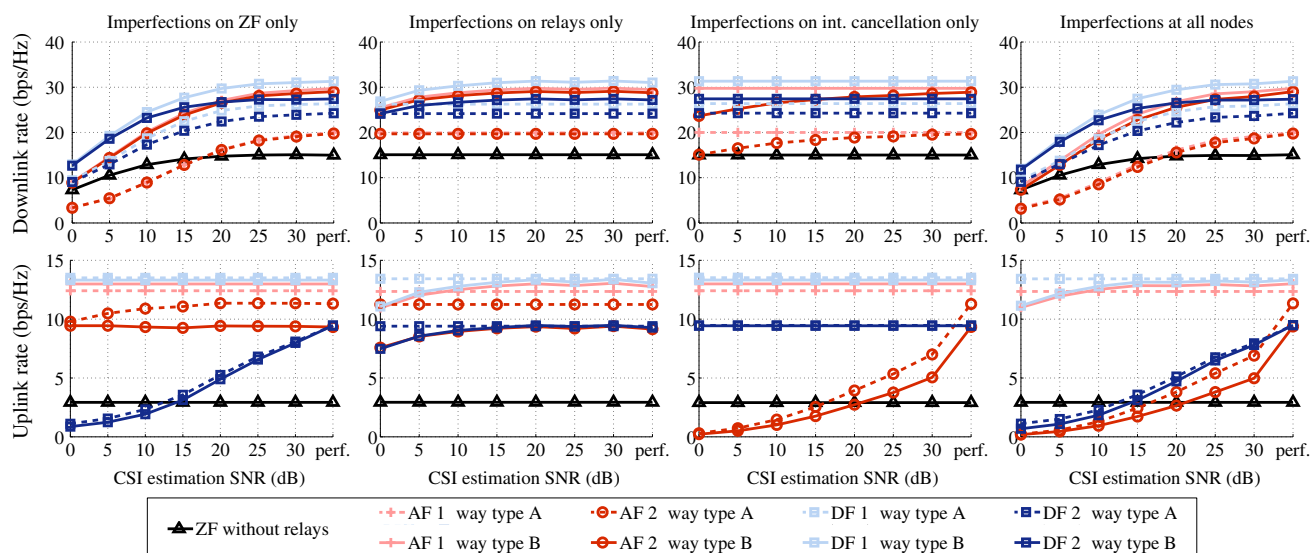


Fig. 8 Influence of imperfect CSI at the different nodes. In the first three columns, only one type of CSI estimation is affected by imperfections, the others are assumed to be perfect. In the column on the right, the estimation SNR is the same at all involved nodes

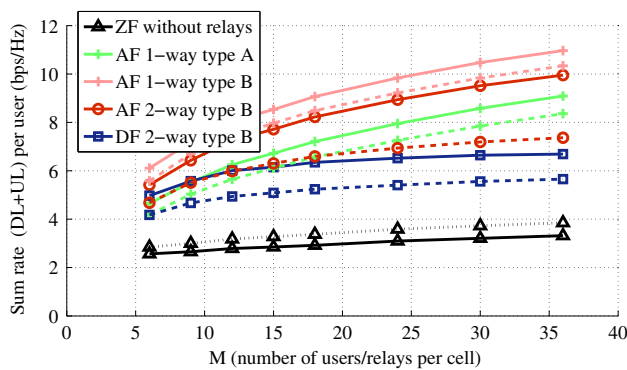


Fig. 9 Increasing number of relays/users with (*dashed*) and without (*solid*) imperfections. The *black dotted line* corresponds to the reference where the BS transmits with the sum power that the BS and the relays would have together

schemes can be attributed to the additional power of the relay nodes, this is not the main source of improvement. The reference can only marginally be increased by adding additional power. This indicates that the reference scenario is mainly interference limited and that an increased transmit power does not lead to considerable improvements. With relaying, on the other hand, the situation can significantly be improved by the relay carpet with the distributed interference mitigation. Interestingly, AF relaying outperforms DF relaying. As more nodes are present in the network, the interference sources become denser and the DF relays have more difficulties to decode their receive signals. The denser network, however, has less impact on AF relaying, which makes it a more suitable choice for the relay carpet. It can also be observed that especially the simple type A AF relays achieve a good performance and the degradation with CSI imperfections are small. As these relays are of very low complexity, more of these relays can be deployed with little costs. The lower rates as compared to the more complex type B relays can thus be recovered by deploying more of them. Also the use of idle MSs as relays can further improve the performance, as a growing network increases the throughput. However, the two-way gain is not as pronounced in networks with high density when the CSI is imperfect. Hence, the possible performance loss of one-way relaying due the multiple channel uses can be recovered by an increased relay density.

7 Conclusions

The relay carpet is a promising approach to stretch the boundaries of cellular networks. By the use of ubiquitous relaying, interference can be reduced and the coverage is more homogeneous. Through the distributed form of

interference management, the spatial degrees of freedom can be better exploited and the frequency reuse factor can be improved towards one. Turning the cellular network into a two-hop network also simplifies CSI estimation at the terminals and enables massive MIMO antenna arrays at the BSs. This approach is not only scalable in terms of the number of involved nodes/antennas, but it is also transparent to the implemented communication technology and can be applied on top of other approaches such as CoMP, heterogeneous networks, or others.

With BSs that can cooperate with each other to cancel other BS interference, two-way AF relaying shows large performance gains and proves to be very efficient for cellular networks. The limited form of BS cooperation introduces only a small overhead because the channels to the relays have a long coherence time. Moreover, no clustering of BSs is required for this form of cooperation; any information that helps to reconstruct and cancel interference is beneficial. If the prelog loss due to the use of multiple channel uses for one transmission is considered, two-way relaying clearly outperforms one-way relaying. On the other hand, the one-way schemes with simple AF relays are very robust with respect to imperfections. The relays can thereby be of very low complexity; especially in FDD, they can be implemented as simple frequency converters. By deploying a large number of them, the throughput of cellular networks can still be enhanced significantly with comparably low costs. If full-duplex relays can be used or when the second hop is for free (e.g. as a secondary link), one-way AF relays can lead to a better performance than two-way relaying.

The sample transmission schemes applied in this paper already show a significant gain as compared to a conventional multi-user MIMO approach. However, the proposed schemes are not optimal in any way. The performance can be further increased, e.g. when the schemes are combined with power control and/or transmit cooperation at the BSs, not least as some user data and CSI is already available at these nodes. Also the possibility to include user cooperation into the proposed network can be beneficial. We thus consider the relay carpet to be a promising option for future cellular networks that can improve their performance by the required factors.

Appendix 1: Derivation of Covariance Matrices

The covariance matrices that are used for the rate calculations in Sect. 3 are derived in the following. For the case of the downlink in AF one-way relaying, the covariance matrix of the desired signal is

$$\begin{aligned} \overrightarrow{\mathbf{K}}_{c,k}^{(\text{sig})} &= \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{F}}_{k,j}^{(c,d)} \mathbf{G}_{d,j} \mathbf{H}_j^{(d,c)} \mathbf{Q}_{c,k}^{(\text{B})} \\ &\quad \cdot \mathbf{Q}_{c,k}^{(\text{B})\text{H}} \mathbf{H}_{j'}^{(d',c)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{F}}_{k,j'}^{(c,d')\text{H}} \end{aligned}$$

and the one of the interference and noise follows as

$$\begin{aligned} \overrightarrow{\mathbf{K}}_{c,k}^{(\text{i+n})} &= \mathbb{E} \left[\overrightarrow{\mathbf{y}}_{c,k} \cdot \overrightarrow{\mathbf{y}}_{c,k}^{\text{H}} \right] - \overrightarrow{\mathbf{K}}_{c,k}^{(\text{sig})} \\ &= \sum_{d \neq c}^C \sum_{j=1}^M \sum_{b=1}^C \sum_{i=1}^K \sum_{b'=1}^C \sum_{i'=1}^K \overline{\mathbf{F}}_{k,i}^{(c,b)} \mathbf{G}_{b,i} \mathbf{H}_i^{(b,d)} \mathbf{Q}_{d,j}^{(\text{B})} \\ &\quad \cdot \mathbf{Q}_{d,j}^{(\text{B})\text{H}} \mathbf{H}_{i'}^{(b',d)\text{H}} \mathbf{G}_{b',i'}^{\text{H}} \overline{\mathbf{F}}_{k,i'}^{(c,b')\text{H}} \\ &\quad + \sum_{j=1}^M \sum_{b=1}^C \sum_{i=1}^K \sum_{b'=1}^C \sum_{i'=1}^K \overline{\mathbf{F}}_{k,i}^{(c,b)} \mathbf{G}_{b,i} \mathbf{H}_i^{(b,c)} \mathbf{Q}_{c,j}^{(\text{B})} \\ &\quad \cdot \mathbf{Q}_{c,j}^{(\text{B})\text{H}} \mathbf{H}_{i'}^{(b',c)\text{H}} \mathbf{G}_{b',i'}^{\text{H}} \overline{\mathbf{F}}_{k,i'}^{(c,b')\text{H}} \\ &\quad + \sigma_n^2 \sum_{b=1}^C \sum_{i=1}^K \overline{\mathbf{F}}_{k,i}^{(c,b)} \mathbf{G}_{b,i} \mathbf{G}_{b,i}^{\text{H}} \overline{\mathbf{F}}_{k,i}^{(c,b)\text{H}} + \sigma_w^2 \mathbf{I}_{N_M}. \end{aligned}$$

In the uplink, they are

$$\begin{aligned} \overleftarrow{\mathbf{K}}_c^{(\text{sig})} &= \sum_{k=1}^M \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{F}_{j,k}^{(d,c)} \mathbf{Q}_{c,k}^{(\text{M})} \\ &\quad \cdot \mathbf{Q}_{c,k}^{(\text{M})\text{H}} \mathbf{F}_{j',k}^{(d',c)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{H}}_{j'}^{(c,d')\text{H}}, \end{aligned}$$

and

$$\begin{aligned} \overleftarrow{\mathbf{K}}_c^{(\text{i+n})} &= \sum_{b \neq c}^C \sum_{k=1}^M \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{F}_{j,k}^{(d,b)} \mathbf{Q}_{b,k}^{(\text{M})} \\ &\quad \cdot \mathbf{Q}_{b,k}^{(\text{M})\text{H}} \mathbf{F}_{j',k}^{(d',b)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{H}}_{j'}^{(c,d')\text{H}} \\ &\quad + \sigma_n^2 \sum_{d=1}^C \sum_{j=1}^K \overline{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{G}_{d,j}^{\text{H}} \overline{\mathbf{H}}_j^{(c,d)\text{H}} + \sigma_w^2 \mathbf{I}_{N_B}. \end{aligned}$$

In the AF two-way case, the downlink covariance matrix of the desired signal at MS (c, k) is given by

$$\begin{aligned} \mathbf{K}_{\text{M},c,k}^{(\text{sig})} &= \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{F}}_{k,j}^{(c,d)} \mathbf{G}_{d,j} \mathbf{H}_j^{(d,c)} \mathbf{Q}_{c,k}^{(\text{B})} \\ &\quad \cdot \mathbf{Q}_{c,k}^{(\text{B})\text{H}} \mathbf{H}_{j'}^{(d',c)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{F}}_{k,j'}^{(c,d')\text{H}}. \end{aligned}$$

For the interference, we distinguish the covariance matrix of self-interference and of the remaining interference plus noise:

$$\begin{aligned} \mathbf{K}_{\text{M},c,k}^{(\text{self})} &= \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{F}}_{k,j}^{(c,d)} \mathbf{G}_{d,j} \mathbf{F}_{j,k}^{(d,c)} \mathbf{Q}_{c,k}^{(\text{M})} \\ &\quad \cdot \mathbf{Q}_{c,k}^{(\text{M})\text{H}} \mathbf{F}_{j',k}^{(d',c)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{F}}_{k,j'}^{(c,d')\text{H}} \\ \mathbf{K}_{\text{M},c,k}^{(\text{i+n})} &= \mathbb{E} \left[\mathbf{y}_{c,k}^{(\text{M})} \cdot \mathbf{y}_{c,k}^{(\text{M})\text{H}} \right] - \mathbf{K}_{\text{M},c,k}^{(\text{sig})} - \mathbf{K}_{\text{M},c,k}^{(\text{self})}. \end{aligned}$$

For the uplink, covariance matrices of the desired signal (jointly from all corresponding MSs), of the BS-interference (including self-interference), and of the remaining interference plus noise follow similarly as

$$\begin{aligned} \mathbf{K}_{\text{B},c}^{(\text{sig})} &= \sum_{k=1}^M \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{F}_{j,k}^{(d,c)} \mathbf{Q}_{c,k}^{(\text{M})} \\ &\quad \cdot \mathbf{Q}_{c,k}^{(\text{M})\text{H}} \mathbf{F}_{j',k}^{(d',c)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{H}}_{j'}^{(c,d')\text{H}}, \end{aligned}$$

$$\begin{aligned} \mathbf{K}_{\text{B},c}^{(\text{BSint})} &= \sum_{b \neq c}^C \sum_{k=1}^M \sum_{d=1}^C \sum_{j=1}^K \sum_{d'=1}^C \sum_{j'=1}^K \overline{\mathbf{H}}_j^{(c,d)} \mathbf{G}_{d,j} \mathbf{H}_j^{(d,b)} \mathbf{Q}_{b,k}^{(\text{B})} \\ &\quad \cdot \mathbf{Q}_{b,k}^{(\text{B})\text{H}} \mathbf{H}_{j'}^{(d',b)\text{H}} \mathbf{G}_{d',j'}^{\text{H}} \overline{\mathbf{H}}_{j'}^{(c,d')\text{H}}, \end{aligned}$$

and

$$\mathbf{K}_{\text{B},c}^{(\text{i+n})} = \mathbb{E} \left[\mathbf{y}_c^{(\text{B})} \cdot \mathbf{y}_c^{(\text{B})\text{H}} \right] - \mathbf{K}_{\text{B},c}^{(\text{sig})} - \mathbf{K}_{\text{B},c}^{(\text{BSint})}.$$

References

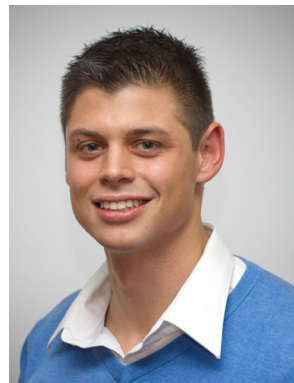
1. R. Rolny, M. Kuhn, and A. Wittneben, The relay carpet: ubiquitous two-way relaying in cooperative cellular networks. In *PIMRC'13*, London, UK, Sep 2013.
2. J. Lee et al., Coordinated multipoint transmission and reception in LTE-advanced systems. *IEEE Communications Magazine*, Vol. 50, No. 11, pp. 44–50, 2012.
3. T. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, *IEEE Transactions on Wireless Communications*, Vol. 9, No. 11, pp. 3590–3600, 2010.
4. T. Nakamura et al., Trends in small cell enhancements in LTE-advanced, *IEEE Communications Magazine*, Vol. 51, No. 2, pp. 98–105, 2013.
5. A. Ghosh et al., Heterogeneous cellular networks: from theory to practice, *IEEE Communications Magazine*, Vol. 50, No. 6, pp. 54–64, 2012.
6. K. Hosseini et al., Massive MIMO and small cells: how to densify heterogeneous networks. In *ICC*, June 2013.
7. W. Choi, and J. Andrews, Downlink performance and capacity of distributed antenna systems in a multicell environment, *IEEE Transactions on Wireless Communications*, Vol. 6, No. 1, pp. 69–73, 2007.
8. H. Viswanathan, and S. Mukherjee, Performance of cellular networks with relays and centralized scheduling, *IEEE Transactions on Wireless Communications*, Vol. 4, No. 5, pp. 2318–2328, 2005.
9. D. Gesbert et al., Multi-cell MIMO cooperative networks: a new look at interference, *IEEE Journal on Selected Areas in Communications*, Vol. 28, No. 9, pp. 1380–1408, 2010.
10. M.K. Karakayali, G.J. Foschini, R.A. Valenzuela, and R. Yates, On the maximum common rate achievable in a coordinated network. In *IEEE International Conference on Communications (ICC)*, June 2006.
11. S.A. Ramprasad, G. Caire, and H.C. Papadopoulos, Cellular and network MIMO architectures: MU-MIMO spectral efficiency and

costs of channel state information. In *Asilomar Conference on Signals, Systems, and Computers*, Nov 2009.

12. A. Lozano, R. Heath, and J. Andrews, Fundamental limits of cooperation, *IEEE Transactions on Information Theory*, Vol. 59, No. 9, pp. 5213–5226, 2013.
13. S. Peters, A. Panah, K. Truong, and R. Heath, Relay architectures for 3GPP LTE-advanced, *EURASIP Journal on Wireless Communications and Networking*, Vol. 2009, No. 1, pp. 11–114, 2009.
14. A. Wittneben, and B. Rankov, Impact of cooperative relays on the capacity of rank-deficient MIMO channels. In *Proceedings of the 12th IST Summit on Mobile and Wireless Communications*, pp. 421–425, June 2003.
15. R. Rolny, J. Wagner, C. Esli, and A. Wittneben, Distributed gain matrix optimization in non-regenerative MIMO relay networks. In *Asilomar Conference on Signals, Systems, and Computers*, Nov 2009.
16. G. Kramer, M. Gastpar, and P. Gupta, Cooperative strategies and capacity theorems for relay networks, *IEEE Transactions on Information Theory*, Vol. 51, No. 9, pp. 3037–3063, 2005.
17. K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, New York, 2009.
18. B. Rankov, and A. Wittneben, Spectral efficient signaling for half-duplex relay channels. In *Asilomar Conference on Signals, Systems, and Computers*, Nov 2005.
19. J. Zhao, M. Kuhn, A. Wittneben, and G. Bauch, Self-interference aided channel estimation in two-way relaying systems. In *IEEE Global Communications Conference (GLOBECOM)*, Nov 2008.
20. D. Tse, and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, New York, 2005.
21. C. Esli, Design and optimization of distributed multiuser cooperative wireless networks, Ph.D. Thesis, ETH Zurich, 2010.
22. B. Rankov, and A. Wittneben, Achievable rate regions for the two-way relay channel. In *IEEE International Symposium on Information Theory*, pp. 1668–1672, July 2006.
23. S. Sun, Y. Ju, and Y. Yamao, Overlay cognitive radio OFDM system for 4G cellular networks, *IEEE on Wireless Communications*, Vol. 20, No. 2, pp. 68–73, 2013.
24. Q. Spencer, A. Swindlehurst, and M. Haardt, Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels, *IEEE Transactions on Signal Processing*, Vol. 52, No. 2, pp. 461–471, 2004.
25. K. Gomadam, V. Cadambe, and S. Jafar, Approaching the capacity of wireless networks through distributed interference alignment. In *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2008.
26. S. Haykin, *Adaptive Filter Theory* (3rd ed.). Prentice-Hall Inc, Upper Saddle River, 1996.
27. T. Weber, A. Sklavos, and M. Meurer, Imperfect channel-state information in MIMO transmission, *IEEE Transactions on Communications*, Vol. 54, No. 3, pp. 543–552, 2006.
28. R. Rolny, M. Kuhn, A. Wittneben, and T. Zasowski, Relaying and base station cooperation: a comparative survey for future cellular networks. In *Asilomar Conference on Signals, Systems, and Computers*, Nov 2012.



processing. His research focuses on cooperative signaling, distributed MIMO, and relaying for cellular and short-range wireless systems and networks. Mr. Rolny has received a Best Paper Award at the IEEE PIMRC 2013.



Tim Rüegg received the M.Sc. degree in Information Technology and Electrical Engineering from the Swiss Federal Institute of Technology (ETH) Zurich, Switzerland, in 2012 with distinction. Since August 2012 he is with the Communication Technology Laboratory at ETH working towards the Ph.D. degree. His research interests are in the field of signal processing for cellular networks.



including space–time coding, MIMO communication, cooperative signal processing, and power line communication.

Marc Kuhn was born in Saarlouis, Germany. He received the Dipl.-Ing. and Dr.-Ing. Degrees in Electrical Engineering from Saarland University, Saarbrücken, Germany, in 1998 and 2002, respectively. Since 2002, he has been with the Communication Technology Laboratory, Swiss Federal Institute of Technology (ETH) Zurich, Zurich, Switzerland, where he is currently a Senior Researcher. His main research interests are in wireless communications



Armin Wittneben received his M.Sc. and Ph.D. degrees in Electrical Engineering from the Technical University of Darmstadt, Germany, in 1983 and 1989 respectively. In 1989 he joined the Central Research Unit ASCOM Tech of ASCOM Ltd., Switzerland, as postdoctoral researcher. In 1992 he became head of the Wireless Technology and Security Department. In 1996 he also served as co-director of ASCOM Tech. In 1997 he received

the TU-Darmstadt. In 1998 he accepted a position as full professor of Telecommunication at the Saarland University, Germany. Since 2002 he has been full professor of Wireless Technology at ETH Zurich, Switzerland and is serving as head of the Communication Technology Laboratory. His research interests include all aspects of wireless communication with special emphasis on communication theory, signal processing, cooperative wireless signaling and protocols, distributed MIMO systems and Ultra-Wideband/Ultra-Low Power communication, imaging and position location.

the Venia Legendi (Habilitation) in Communication Technology from